**Technical report and analysis of speak week 2021 students’ feedbacks**

**Outline**

* Introduction
* Aims
* Exploratory Data Analysis
* Statistical Summary
* Missing Value Analysis
* Missing Value Heatmap
* Missing data Histogram plot
* Features Exploration
* Data Cleaning
* Exploration of cleaned dataframe
* Statistical Analysis and Visualization of column Q1-Q4
* Exploration of Question 1&4 i.e. (Q1) and (Q2)
* Exploration of Question 2&4 i.e. (Q2) and (Q4)
* Visualisation Grouped bar chart Q1-Q4
* Sentiment Analysis of Q5 (All Comments) using wordcloud and nltk modules
* Statistical Analysis of cleaned and uncleaned Q5 column
* Conclusions

**Introduction**

This report is based on the exploratory data analysis of student’s feedbacks collected during Speak Week event organised in 2021 by University of Bournemouth Student Union which was designed to get students opinions about returning to the university campus for face-to-face lectures and academic activities which had been previously impacted by covid 19 pandemic i.e., social distancing and other rules designed to mitigate the spread of corona virus. Five Questions were asked in the survey, four quantitative (multiple choice) questions i.e., question 1 – 4 and one qualitative question, free text i.e., question 5. The feedbacks were curated into 5 columns, 346 rows in the excel spreadsheet.

**Aims**

The aim of this study is to analyse student’s response from the Speak Week organised by the SUBU, draw meaningful insights which could be transformed into decisions, guidelines and policies that would protect students, teaching and non-teaching members of the Bournemouth university and the community during the covid 19 pandemic. The goal is to develop a model that keeps everyone safe while student pursue their academic ambitions on the university campus.

**Exploratory Data Analysis**

The dataset was loaded in the jupyter notebook as dataframe using Panda’s library in python. The dataset has five categorical columns and 346 rows, with some missing values displayed as NaN in the 5th Column i.e., Question 5(Please tell us more about your answer to the previous question). Question 1-4 has 346 non – null entries and classified as object datatype. Question 5(5th column) has 262 non-null entries, 84 null entries i.e. (NaN) and classified as object data type.

**Statistical Summary of column 1-4**

The statistical summary of question 1 showed that column 1 has 346 entries, 6 unique age range; under 21, 21-24, 25-29, 30-39, 40-49, 50+, and the top count as age under 21 with frequency of 151, which means that 151 under 21 students responded to the survey i.e., 43.64%

The statistical summary of questions 2 showed that column 2 has 346 entries, 5 unique faculties: Faculty of Science and Technology, Faculty of Media and Communication, The Business School/Faculty of Management, Faculty of Health, and Social Sciences, NCCA. The highest response in terms of faculty level was recorded from students from Faculty of Science and Technology, 117 students from Faculty of science and Technology responded to the survey i.e., 33.82%

The statistical summary of questions 3 showed that column 3 has 346 entries, 7 unique year of study/study level: 1st Year (including Foundation year), 3rd Year or Final Year, 2nd Year, Post-Graduate (non-doctorate), PhD or equivalent, I can't say which year due to pandamic my data collection is still withholding, Placement Year. The highest response based on study year was recorded from 1st Year (including Foundation year), students. 122 1st Year (including Foundation year) students responded to the survey i.e., 35.26%

The statistical summary of questions 4 showed that column 4 has 346 entries, 4 unique responses; Very comfortable, I can’t wait to get back, Not at all comfortable, I have serious concerns, Somewhat comfortable, but I have some concerns, Comfortable, as long as everyone respects physical distancing rules, etc. The top response recorded was Very comfortable, I can’t wait to get back, 138 students responded as “Very comfortable, I can’t wait to get back” i.e. 39.88%

**Missing Value Analysis**

Missing values are the empty, null entries i.e., NaN in the question 5 (5th column of the dataset). There are 84 NaN values i.e., 24.28%, this could mean that 84 students didn’t provide a response to the question 5 (Please tell us more about your answer to the previous question) or there was an error of omission during data entry task.

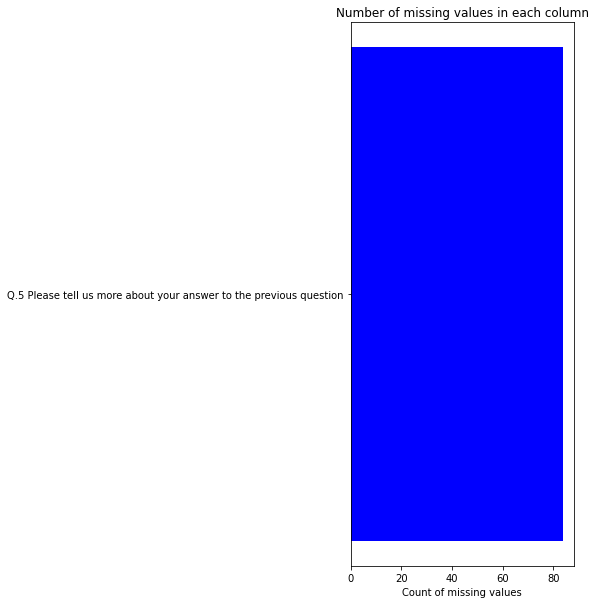
Figure 1 Missing value heatmap

**Shape

Description automatically generated with medium confidence**

**Missing data Histogram plot**

Figure 2 Missing data histogram plot

****

**Features Exploration**

Subsequently for better analytic result the question 1-5 columns headings were renamed as Q1, Q2, Q3, Q4, Q5 using the df.rename function in python. The columns Q1-Q5 were explored, and unique features viewed using the df.unique() function in python.

**Data Cleaning**

A function in python was created to clean the dataset by replacing the null entries i.e., the NaN in column Q5 with unknown, the dataset was furthermore explored to confirm changes.

Statistical Analysis and Visualization of Columns Q1-Q4

Table 1 Response rate Analysis of Q1/Age range

|  |  |  |
| --- | --- | --- |
| Age range | Counts | Percentage |
| Under 21 | 151 | 43.64% |
| 21- 24 | 131 | 37.86% |
| 25-29 | 27 | 7.8% |
| 30-39 | 26 | 7.51% |
| 40-49 | 7 | 2.02% |
| 50+ | 4 | 1.16% |

Figure 3 count plot of Q1

**Chart, bar chart

Description automatically generated**

Table 1 and fig 3 shows that students age under 21 and 21-24 participated actively in the survey, however other age groups had low participation.

Table 2 Response Rate Analysis of Q2 column/Faculty

|  |  |  |
| --- | --- | --- |
| **Faculty** | Counts | Percentage |
| Faculty of Science and Technology | 117 | 33.82% |
| Faculty of Media and Communication | 86 | 24.86% |
| The Business School/Faculty of Management | 74 | 21.39% |
| Faculty of Health and Social Sciences | 68 | 19.65% |
| NCCA | 1 | 0.29% |

Figure 4 Countplot of Q2

Chart, bar chart

Description automatically generated

Table 2 and fig 4 shows that students from Faculty of Science and Technology, Faculty of Media and Communication, The Business School/Faculty of Management, Faculty of Health and Social Sciences participated actively in the survey, however NCCA had low participation.

Table 3 Response Rate Analysis of Q3/ study year

|  |  |  |
| --- | --- | --- |
| Study Year | Counts | Percentage |
| 1st Year (including Foundation year) | 122 | 35.26% |
| 2nd Year | 83 | 23.99% |
| 3rd Year or Final Year | 69 | 19.94% |
| Post-Graduate (non-doctorate) | 48 | 13.87% |
| Placement Year | 12 | 3.47% |
| PhD or equivalent | 11 | 3.17% |
| I can't say which year due to pandamic my data collection is still withholding | 1 | 0.29% |

Figure 5 Countplot of Q3

**Chart, bar chart

Description automatically generated**

Table 3 and fig 5 shows that students from 1st year, 2nd year and 3rd year participated actively in the survey, however postgraduate, PhD and placement year students had low participation.

Table 4 Response Rate Analysis of Q4 column/ Response

|  |  |  |
| --- | --- | --- |
| Response | Counts | Percentage |
| Very comfortable, I can’t wait to get back | 138 | 39.88% |
| Comfortable, as long as everyone respects physical distancing rules, etc. | 122 | 35.26% |
| Somewhat comfortable, but I have some concerns | 62 | 17.92% |
| Not at all comfortable, I have serious concerns | 24 | 6.94% |

Figure 6 Countplot of Q4

**Chart, bar chart

Description automatically generated**

Table 4 and fig 6 shows that 93.06 students are very comfortable with getting back to the university campus( this includes comfortable without condition, comfortable with condition and comfortable with some concerns). 6.94% of students are not comfortable with getting back and have serious concerns.

**Exploration of Question 1&4 i.e. (Q1) and (Q4)**

Figure 7 bar plot of Q1 AND Q4

A screenshot of a computer

Description automatically generated with medium confidence

**Potential link between a respondent’s age (Q1) and the responses to Question 4(Q4)**

* Student age Under 21 records the highest count for “very comfortable, I can’t wait to get back.”
* Student age under 21 records the highest count for “Comfortable, as long as everyone respects physical distancing rules etc”
* Student age under 21 records the second highest count for “somewhat comfortable, but I have some concerns”.
* Student age under 21 records the second highest count for “Not at all comfortable, I have serious concerns”
* Student age 21-24 records the second highest count for” very comfortable, I can’t wait to get back.”
* Student age 21-24 records the second highest count for “Comfortable, as long as everyone respects physical distancing rules etc”
* Student age 21-24 records the highest count for “somewhat comfortable, but I have some concerns”
* Student age 21-24 records the highest count for “Not at all comfortable, I have serious concerns”
* Student age 25-29 records low response to the question Q4, and zero record for “ Not at all comfortable, I have serious concerns”
* Student age 30-39 records low response to the question Q4
* Student age 40-49 records low response to the question Q4
* Student age 50+ records low response to the question Q4 and zero for “somewhat comfortable, but I have some concerns” and “Not at all comfortable, I have serious concerns”

In summary it is evident that response to question Q4 decreases across the age group in ascending order. Student age under 21 years are interested in getting back to the university campus, if protective procedures were in place to mitigate the spread of covid 19, as with any population group, some percentages are still worried about getting infected. Student age under 21, 21-24 have stronger immune system and are excited to return to their friends and the university lifestyle.

Furthermore, student age 21-24 are also interested in returning to the university campus as long as the spread of covid-19 is contained while certain rules are in place, some are also worried about getting infected. However, student age 25-29, 30-39, 40-49, 50+ are less interested in returning to the university campus, this is expected as they might have partners, kids that they are worried about, most especially student age 50+ may have underlying health conditions, weaker immune system and are more vulnerable to covid 19 virus.

**Exploration of Question 2&4 i.e. (Q2) and (Q4)**

Figure 8 bar plot of Q2 AND Q4

Chart, bar chart

Description automatically generated

**Potential link between respondent’s faculty (Q.2) and responses to Question 4**

* NCCA students records the lowest response to Question 4, just a few percentages responded to “Very comfortable, I can’t wait to get back” which is the lowest count overall amongst other faculties.
* Faculty of science and technology student records the highest for “Very comfortable, I can’t wait to get back”.
* Faculty of media and communication records the second highest for “Very comfortable, I can’t wait to get back”.
* The business school/faculty of management records the third highest for “Very comfortable, I can’t wait to get back”.
* The Faculty of Health and social sciences records the fourth highest for “Very comfortable, I can’t wait to get back”
* The Faculty of Health and social sciences records the highest for “Comfortable as long as everyone respect social distancing rules, etc”
* Faculty of science and technology student records the second highest for “Comfortable as long as everyone respect social distancing rules, etc”
* The business school/faculty of management records the third highest for “Comfortable as long as everyone respect social distancing rules, etc”
* Faculty of media and communication records the fourth highest for “Comfortable as long as everyone respect social distancing rules, etc”.
* Faculty of science and technology student records the highest for “somewhat comfortable, but I have some concerns”

In summary Faculty of science and technology students are excited about returning to the campus, although they have some concerns but if everyone respect social distancing and other covid 19 protocols they are willing to go back to the university environment, some few percentages are not comfortable with returning to the university campus. NCCA student recorded very low response from the survey, probably because they can collaborate and work effectively remotely as evident in the technology industry.

All other faculties are willing to go back to the university campus, this can be measured from the group bar plot above in fig 8, although they have some concerns but as long as covid 19 protocols are put in place, they would return to the university, some outliers are expected in the data as evident in the response distribution amongst the faculties and age group.

**Exploration of Question 1&2 i.e. (Q1) and (Q2)**

Figure 9 bar plot of Q1 and Q2

Chart

Description automatically generated

The group bar plot above indicates that most students age under 21 and 21-24 , are in Faculty of Science and Technology, Faculty of media and communication and The business School/ faculty of management. Age 25-29, 30-39, 40-49 and 50+ represent the minority groups across the faculties.

**Exploration of Question 1&3 i.e. (Q1) and (Q3)**

Figure 10 bar plot of Q1 and Q3

Chart

Description automatically generated

The group bar plot above in fig 10 indicates that most age under 21, 22-24 are in 1St Year( Including foundation year), 2nd year and 3rd Year or final year. Students age 50+ are mainly in PhD or equivalent, unknown status due to the pandemic. Few students age 40-49 are in 1st year(Including Foundation Year), 2nd year and PhD or equivalent. Few students age 30-39 are mostly Post graduate students(non-doctorate), while some others are in 1st year, 2nd year, and very few are in PhD or equivalent, and Placement year. Most Age 25-29 are Post graduate students, while few others are in 1st year, 3rd year, Phd or equivalent, and very few are in 2nd year. Very few age 21-24 students are in PhD or equivalent, some are in placement year, few are post graduate students.

**Exploration of Question 2&3 i.e. (Q2) and (Q3)**

Figure 11 bar plot of Q2 AND Q3

**Chart, bar chart

Description automatically generated**

The group bar plot above shows the relationship between question 2 and 3, Faculty of science and technology has the heighest 1st year (Includng fundation year), and 2nd year studeents response. NCCA records the lowest response from 2nd year students.

**Exploration of Question 3&4 i.e. (Q3) and (Q4)**

Figure 12 bar plot of Q3 and Q4

**Chart, bar chart

Description automatically generated**

The group bar plot above shows that 1st year students are very comfortable and can’t wait to get back to the university campus, although there are concerns for safety, however they would return to campus as long as everyone respect covid rules. 2nd year and 3rd year students shares similar opinion.

**Sentiment Analysis of Q5 (All Comments) using wordcloud and nltk modules**

Figure 13 word cloud plot

**Text

Description automatically generated**

Figure 14 wordcloud plot

**Text

Description automatically generated**

Figure 15 wordcloud plot

From the word cloud generated using natural language processing(NLP) module in python, the main theme of the Q5 free text can be seen in bold fonts such as people, would, want campus, student, comfortable, course, return, able, get back, feel, social distancing, mask, excited. The majority of this wordcloud text indicates positive emotions of the students. In summary high percentage of student are excited and wants to get back to the university as long as covid protocols are put in place to keep everyone safe,

**Statistical summary of Column Q5 uncleaned raw data**

The statistical summary of questions 5 showed that column 5 has 262 entries, 84 missing values, 248 unique comments, the top comment as “I’m eager to return to the classroom teaching environment and believe that’s the majorities opinion so I would hope everyone would respect the rules in place to allow that to happen and continued doing so” which was made in the survey i.e., 1.15%

**Statistical summary of Column Q5 cleaned data**

The statistical summary of questions 5 showed that column 5 has 346 entries, 84 unknown, 249 unique comments, the top comment as “unknown” which was made in the survey i.e., 24.28% of students didn’t provide more information in Question 5, or this was an error of mission during data entry task.

**Conclusion**

93.06% of Bournemouth University student are excited(including comfortable, comfortable on condition, comfortable with concerns) to return to the university campus for face-to-face academic and extracurricular activities most especially age under 21 and age 21-24, there are some concerns about safety amongst all age groups that participated in the survey, age groups 25-29, 30-39, 40-49 and 50+ are not really comfortable with returning to the university, they would have a positive attitude if protocols are put in place to keep everyone safe such as social distancing rules, face mask, regular handwashing, regular testing, walk in vaccination centres and other government guidelines to mitigate the spread of covid 19 on the university campus.

Age groups 25-29, 30-39, 40-39, and 50+ seems not to participate actively in the survey, I would encourage the SUBU to form policy of inclusion which carries everyone along in all the activities of the union, a good example is to encourage the age range to participate actively as officials of the student union. Also, they are the vulnerable group with weaker immune systems, regular medical checks up, testing, vaccination should be organized in collaboration with the NHS to help keep them safe and healthy.

The students from the department of National Centre for Computer Animation (NCCA) should be encouraged to participate actively in future surveys, this is to provide a complete list of student’s feedbacks, and to assist the SUBU in forming a wholistic approach in solving student academic and welfare problems.